DOCUMENT RESUME

ED 268 935 HE 019 303

AUTHOR Coughlin, Lorence D.; Patel, Vimla L.
TITLE Text Comprehension and Expertise in the Domain of Medicine.
SPONS AGENCY Josiah Macy, Jr. Foundation, New York, N.Y.; Natural Sciences and Engineering Research Council, Ottawa (Ontario).
PUB DATE [86]
NOTE 53p.; Paper presented at the Annual Meeting of the American Educational Research Association (70th, San Francisco, CA, April 16-20, 1986).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Clinical Diagnosis; Cognitive Processes; *Difficulty Level; Discourse Analysis; Encoding (Psychology); Higher Education; *Medical Case Histories; *Medical Students; *Physicians; *Reading Comprehension

ABSTRACT

The effects of the randomization of information in clinical cases on the comprehension of 16 physicians and 16 medical students were studied, using structured and unstructured forms of two clinical cases. Attention was directed to the effects of case difficulty and text structure on representations of cases in written protocols, ability to access the relevant and critical information in the cases, and ability to generate diagnoses. The research paradigm was drawn from cognitive science and the analysis techniques were based on the techniques of discourse analysis. when viewed in terms of assessing "relevant" information, major findings indicate that both experts and novices in medicine are affected by the structural form of clinical cases and that this effect interacts with the difficulty level of the case. These effects were not present when viewed in terms of access to "critical" information. The comprehension of information in an easier clinical case was found to replicate the findings of studies in other areas of expertise. Difficult cases were found to be processed in a different manner. Experts were differentiated from novices by their ability to selectively encode relevant and critical case information and by the quality of their diagnoses. (Author/SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

HE 019 303

ED268935

TEXT COMPREHENSION AND EXPERTISE IN THE DOMAIN OF MEDICINE

LORENCE D. COUGHLIN

VIMLA L. PATEL

CENTRE FOR MEDICAL EDUCATION
McGILL UNIVERSITY
MONTREAL, CANADA

HE 019 303

Reprints:

Lorence D. Coughlin
Centre for Medical Education
McGill University
3655 Drummond Street - Room 529
Montreal, Quebec. H3G 1Y6

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Lorence D. Coughlin

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ACKNOWLEDGEMENTS

We wish to thank Drs. H.M. Scott and H. HoPingKong for their assistance with the medical aspects of this study. We acknowledge the support of the Internists and Medical Students who generously volunteered their time. This work was supported in part by Grants from the Natural Science and Engineering Council of Canada (# 2598) and Josiah Macy Jr. Foundation to Vimla Patel.

ABSTRACT

The effects of the randomization of information in clinical cases on the comprehension of physicians and medical students were studied through the analysis of written recall protocols of structured and unstructured forms of two clinical cases. The research paradigm in this study was drawn from cognitive science and the analysis techniques were based on the techniques of discourse analysis.

The major findings of the study indicate that, when viewed in terms of accessing 'relevant' information, both experts and novices in medicine are affected by the structural form of clinical cases and that this effect interacts with the difficulty level of the case. These effects were not present when viewed in terms of access to 'critical' information.

The comprehension of information in an easier clinical case was found to replicate the findings of studies in other areas of expertise. Difficult cases were found to be processed in a different manner. Experts were differentiated from novices by their ability to selectively encode relevant and critical case information and by the quality of their diagnoses.

Literature Review

The study of expert-novice differences, based on the paradigm developed in the classic studies of chess by deGroot (1965) and Chase and Simon (1973a), has led to the characterization of the comprehension and problem solving abilities of subjects at varying levels of expertise in many different domains. Basic to this research paradigm is the inclusion of stimulus materials with a structure which has been randomized in order to view the abilities of experts and novices in both 'typical' and 'atypical' situations.

The results of the studies on chess have shown that both experts and novices are severly limited in the number of pieces of information that can be stored in short-term memory but experts "chunk" the pieces in normal, structured chess board configurations. In other words, when the chess boards presented are meaningful mid-game configurations, the experts can reconstruct four or five chunks of information, with each chunk containing four or five individual chess men. When the boards are randomly arranged, the experts cannot chunk the information, although they do try, and, consequently, they do no better than the novices, placing only four or five men correctly.

In hypothesizing about the chess master's ability, Chase and Simon (1973b) believed that experts recognized and internally represented the relevant patterns from the chess board, a conceptual task. Novices, on the other hand, tended to perform a purely perceptual task, neither recognizing nor differentiating relevant from irrelevant patterns.

Similar results have been found in many other areas of expertise. In a study

of expertise in electronics (Egan and Schwartz, 1979), experts were much more effective than novices at recalling 'typical' schematic drawings of circuit boards. The superior performance of experts was attributed to their ability to organize their representations of the boards on the basis of function. Novices, on the other hand, performed the task on the basis of the spatial proximity of the parts of the boards.

In the comprehension of physics problem, Chi, Feltovich and Glaser (1981) and Larkin, McDermott, Simon and Simon (1980) found that there were distinct differences in the way in which physics problems were catagorized and represented by expert and novice subjects. Those features defined as relevant to solving the problem by novices were different from those identified as relevant by experts, with essentially no overlap of features mentioned by the two groups. Novices catagorized physics problems based on surface-level aspects such as the objects involved (e.g., springs, blocks, pulleys), while the expert subjects catagorized problems according to principles or abstract solution procedures (e.g., conservation of energy, Newton's Second Law).

In the investigation of expert-novice differences in medicine (e.g., Muzzin, Norman, Feightner, Tugwell and Guyatt, 1982; Norman, Feightner, Jacoby and Campbell, 1979) many of the methods of analysis used have focused on the subjects' ability to reconstruct the explicit text-base (Kintsch and van Dijk, 1978) in a bottom-up or data-limited (Bobrow and Norman, 1975) approach to comprehension of complex verbal data. Comprehension, however, has been hypothesized to involve interactive bottom-up and top-down processing (Bartlett, 1932; Danks and Glucksburg, 1980). This implies that both the text-base and the comprehender's knowledge-base are involved in comprehension and both must be

considered if the true nature of the differences in expertise are to be accessed.

This problem is evident in the numerous studies in medicine which have resulted in the anomolous finding of no differences between the comprehension of experts and novices. Groen and Patel (1985) have questioned the feasibility of transposing the techniques used in the analysis of the well-structured, non-verbal tasks in Chase and Simon's chess experiments directly to the ill-structured, complex verbal task of medical problem-solving.

The counter-intuitive findings of many of the studies on expertise in medicine led Patel and Frederiksen (1984a) to perform a reanalysis of some of the data from the Muzzin et al. (1982) study using a technique of detailed semantic analysis as developed by Frederiksen (1975) and adapted to medicine by Patel and Frederiksen (1984a,b).

The results of this reanalysis showed significant differences by level of expertise, both within and across typical and atypical stimulus conditions and by the response type. The Patel and Frederiksen method of analysis delineates two basic types of responses in recalling clinical case information. These response types are **'recall',** reconstuctions of portions of the case which are drawn directly from the original text base, and **'inference',** transformations performed on the original text base, which are the result of the interaction between the text and the subjects' world knowledge, specific or general. These transformations, which involve higher level processes than recall (Collins, Brown and Larkin, 1980), are defined by a set of specific rules which are detailed in the Protocol Analysis subsection below.

In dealing with the comprehension of complex verbal stimuli, an accepted factor determining the likelihood of any given unit of information being recalled

is its relevance or salience (Johnson, 1970; Pichert and Anderson, 1977; Borland and Flammer, 1985). In a clinical problem-solving situation, this aspect is interpreted as refering to the determination of relevant or critical units of information in relation to reaching an accurate diagnosis (Patel, Groen and Frederiksen, 1986; Norman, Tugwell, Feightner, Muzzin and Jacoby, 1985).

The development of randomized or 'atypical' stimuli in the domain of medicine has proven to be somewhat problematic (Coughlin and Patel, 1985). The typical-atypical differentiation has been defined in two very different manners. One approach is clinically based, defined in terms of the difficulty (e.g., easy-hard) or the rate of occurance of the particular case (e.g., common-rare) as in the study by Muzzin and her colleagues (1982). The other approach is based on manipulations of the case information in such a way that the stimulus is constructed so as to not be representative of any specific disease (e.g., Claessen and Boshuizen, 1985), thus making it "atypical".

In the chess experiments identical and directly comparable items (the individual chess men) were used in both typical and atypical stimuli, there is no such relationship in any of the medical studies. Each case used is totally unique in its content. Therefore, caution must be exercised as any comparisons or conclusions based on the original chess studies may not be totally appropriate.

Patel, et al., (1986) reanalyzed the Muzzin data a second time using the same methodolgy as the prior study (Patel and Frederiksen, 1984a), but included the additional factor of determining the relevancy of the information in the protocols that would generate an accurate diagnosis. Again, highly significant results were found for level of expertise and typicality of the text. The effect of the relevancy of information was also found to be significant. In this study,

the authors also reported new data of their own from a study which they conducted in order to further support their previous findings. They did not, however, use typical and atypical texts because, according to them, they "were not able to come to a clear criterion for atypicality".

This new research did not use clinical cases, but rather, utilized general texts in medicine derived from a general medical textbook (Harrison, 1980). The results showed that expert and novice physicians represent medical text information differently. The physicians were able to selectively encode the relevant information much more efficiently than the novices, who processed relevant and nonrelevant information equally. The experts also made greater use of their prior knowledge, making more inferences than novices, who have been shown to be ineffective in utilizing what prior knowledge they may possess (Bordage and Allen, 1982; Patel, HoPingKong and Mark, 1984).

The purpose of this study is to show what effect the randomization of text information in a written clinical case has on novices and experts in medicine. The approach to be used is derived from the research paradigm developed by Chase and Simon (1973a) in their study of expert-novice differences in chess.

As Reitman (1976) and Patel and Groen (1986) have suggested, the analysis techniques used by Chase and Simon may be inappropriate for more complex conceptual tasks, and attempts to use this methodology in medicine have resulted in mixed success. The analysis of the complex verbal task of clinical case comprehension requires a more detailed method of analysis. Patel and Frederiksen (1984a) have successfully adapted and used the techniques of discourse analysis in the field of medicine to show differences in comprehension as are reflected in the representation of clinical case information.

The suggestion that higher level (Gentner, 1976), or more important (Johnson, 1982) information is more likely to be remembered, particularly by those with more technical knowledge (Krulee, Fairweather and Berquist, 1979) has been intimated in several studies of expertise in medicine (Muzzin et al., 1982; Claessen and Boshuizen, 1985) and has been shown by Patel, Groen and Frederiksen (1986) through the use of measures of "relevancy of information" and by Norman et al. (1985) using "critical information".

The random ordering of case information within a clinical case is an attempt at an adaption of the Chase and Simon methodology, but also included in the research design is the 'typical-atypical' or 'easy-hard' distinction used by several researchers in medical education. To this end two clinical cases were developed which differed in terms of their difficulty to diagnose, one easier case (Acute Bacterial Endocarditis) and one harder case (Temporal Arteritis). Two forms of each case were constructed, a typical ordered structure and a random unstructured form. These were given to both a group of expert physicians and a group of novice medical students. The effects of case difficulty and text structure on the groups was viewed in terms of their representations of the cases in written protocols, their ability to access the relevant and critical information in the cases and their ability to generate the diagnoses.

Experimental Methods

Subjects

Thirty-two subjects were selected at random from two levels of medical expertise, with sixteen subjects from each level. The 'novice' group of subjects consisted of students early in their second year of the medical program at McGill University. The 'expert' group was comprised of Family Medicine practitioners from various teaching hospitals associated with McGill University's Faculty of Medicine.

Materials

Cases: Two clinical cases, developed from actual patients, were used as the stimulus texts for the experiment. They included a case of Acute Bacterial Endocarditis (ABE) and a case of temporal arteritis (TA). The ABE case concerns a 27 year old, unemployed male who arrives at the emergency room complaining of fever and chills. The basic diagnosis of endocarditis is fairly explicit in the text in that the symptoms of toxicity, fever, prostration, the transient loss of vision, shortness of breath, collapsing pulse, and the 2/6 early diastolic murmur in the aortic area are all the result of the vegetation and deterioration of one of the heart valve leaflets. This is usually caused by a bacterial infection. The source of this bacterial infection is found in the fact that the patient is an intravenous drug user. This fact is not explicit in the text, but rather is implied by the patient's claim that he had been bitten by a cat, while the only puncture wounds were found on the inside of the patient's left forearm. The case is deemed as acute because the chills, rigors and the persistant, high fever are of only four

days duration, there is no splenomegaly and the apex beat is normal, all of which denote a sudden onset.

The TA case concerns a 69 year old housewife suffering from the inflammation of one or more of her arteries. Temporal arteritis occurs in elderly persons and may be manifested by constitutional symptoms, severe temporal headaches, and occular symptoms including sudden loss of vision in one eye, all of which are present in this case. This case also includes the important laboratory finding of a greatly increased erythrocyte sedimentation rate (ESR) which is indicative of temporal arteritis.

In a pilot study using the original clinical cases, students and physicians rated the cases in terms of the comprehensibility of the texts and the level of difficulty of the cases on a 5-point scale (1 being low and 5 high). The two cases were found to be roughly equivalent in terms of their comprehensibility (ABE-4.2, TA-4.0), but the TA case was rated as being more difficult than was the ABE case (ABE- 3.0, TA-4.5).

The text on the ABE case consisted of 234 words in 26 segments with a total of 88 propositions. This gave the text a propositional density (i.e., the average number of propositions per segment) of 3.38. The text of the TA case had 187 words in 23 segments with 68 propositions, for a propositional density of 2.96.

Structural Form: The text of each case had two forms, a structured form and an unstructured form. In the structured form the cases were arranged in the typical order of clinical cases familiar to both novices and experts. This structured form follows that used in medical case charts, which is organized in

the following manner:

1) Patient's personal data
2) Medical history
3) Physical findings
4) Laboratory findings

The unstructured forms of the cases contained exactly the same information as the structured forms, but they were randomly ordered.

The basic unit for the randomization of the cases was the segment. The segments of the original, ordered texts were randomly reordered with the only limitations being that segments that were directly related to each other, those connected by the words 'but' or 'when' for example, were retained as ordered pairs and that those segments which contained the most relevant information to the case were separated from each other by at least two other segments. This latter stipulation was utilized in order to avoid substantiating a diagnosis with only a small portion of the text. Segmentation and relevancy are explained in detail in the analysis section below.

Procedure

Both groups of subjects were tested on both clinical cases. Each subject was given one clinical case in the stuctured form and the other clinical case in the unstructured form. This resulted in two subgroups of eight subjects at each level of expertise, with each subgroup having the two cases in different structural forms. The order of presentation of the cases and structuredness were counterbalanced to control for any possible order effects.

An instruction cover sheet was provided with each set of texts. The subjects were instructed to read the first case with a view to recalling the information in it. They were given three and a half to four minutes to read the text (as determined in a pilot study). They were then to turn to the next page and answer the question (Recall the case as accurately as possible and provide a diagnosis.). They were not to refer back to the text at any time once they had read it. They then repeated the procedure with the second case. In addition, they were explicitly instructed to write in longhand and not use short note form. Verbal instructions were given to stress the importance of reading the instructions carefully and to reinforce the fact that subjects were not to refer back to the texts at any time once they had finished reading.

Methods of Analysis

The methods of analysis used in this study are complex and highly detailed, resulting in a microlevel analysis of the recall protocols of the subjects. The methodology is multi-leveled involving five levels of analysis as well as the overall statistical analysis and tabulation of the diagnoses. These multiple levels are:

1) Segmentation
2) Propositional Analysis
3) Protocol Analysis
4) Analysis of Relevant Information
5) Analysis of Critical Information
6) Statistical Analysis
7) Tabulation of Diagnoses

Each level of analysis is described in detail below with examples, where appropriate.

1) Segmentation: At the first level of analysis, all texts, both stimulus materials and subjects' protocols, were segmented using a methodology delineated by Dillinger (1984). This system is based on Winograd's system of clausal analysis (Winograd, 1972,1983) and derived from the systemic grammar of Halliday (1967). This system classifies discourse into major and secondary syntactic units or clauses. As the only secondary clauses treated as separate segments are bound adjuncts, only major clauses and bound adjuncts are differentiated for our purposes. These clausal units facilitate the identification and further analysis of the various parts of the text.

In this system of clausal analysis the general rule, with few exceptions, is that there are the same number of segments in a text as there are finite verbs. That is to say, a clause with a finite (tensed, conjugated) main verb is a separate segment. Clauses with non-finite (infinite, participial) main verbs are viewed as being a part of some other segment rather than as being separate segments on their own.

Bound adjuncts (BAJT) are clauses which are linked to other clauses by explicit binders. The general rule, described previously, continues to apply: i.e., only if the bound adjunct contains a finite verb is it treated as a separate segment. Examples of these rules are given below, the verbs are underlined and slashes (/) indicate the segment boundries.

1. There <u>was</u> no splenomegaly/
1 segment-1 clause

2. She also complained of difficulty getting out of bed in the morning/
1 segment-2 clauses

3. He also complained of some shortness of breath/ when he tried to climb the two flights of stairs in his apartment.
2 segments (1 BAJT)-3 clauses

2) Propositional analysis: After the stimulus texts were segmented, they were propositionally analyzed. An example of this type of analysis is presented in Appendix A. A propositional system is an alternative representation of a semantic network of concepts and relations among concepts. Several propositional systems have been proposed (Kintsch, 1974; Frederiksen, 1975; Norman and Rumelhart, 1975). Frederiksen's (1975,1984) method was used in this study as it has been found to be more fine-grained and comprehensive than the other available systems (Tierney and Mosenthal, 1982) and it has been adapted specifically for the domain of medicine (Patel and Frederiksen, 1984a).

In this system of semantic, conceptual representation each proposition is formally an n-position relation which consists of a label, a head element or predicate and a list of labelled arguments. The label consists of the segment number and the proposition number within the segment separated by a period (e.g., 5.2 refers to segment 5, proposition 2). A predicate can be an action (e.g., climb), an object (e.g., stairs) or a relation which connects propositions (e.g., EQUIV:TEM- a temporal relation of events happening at the same time). The arguments, separated by commas, can be identifying relations (e.g., ATT:- an attribute) or case relations (e.g., AGT:- the agent of a resultive action). Those passages which were used as examples for segmentation are propositionally analyzed below.

1. THERE WAS NO SPLENOMEGALY

1.1 SPLENOMEGALY =NUM:NULL(NO),TNS:PAST;

2.SHE ALSO COMPLAINED OF DIFFICULTY GETTING OUT OF BED IN THE MORNING

2.1 COMPLAIN AGT:SHE,THM:2.2,=TNS:PAST;
2.2 GET RSLT:2.3,=ATT:DIFFICULTY, TEM:(IN)MORNING,ASPCT:CONT;
2.3 (SHE) LOC:(OUT OF)BED;

3. HE ALSO COMPLAINED OF SOME SHORTNESS OF BREATH

3.1 COMPLAIN AGT:HE,THM:3.2,=TEM:___, TNS:PAST;
3.2 BREATH =ATT:SHORTNESS,DEG:SOME;

4. WHEN HE TRIED TO CLIMB THE TWO FLIGHTS OF STAIRS IN HIS APARTMENT

4.1 TRY PAT:HE,GOAL:(TO)4.2,=TEM:4.6, TNS:PAST;
4.2 CLIMB PAT:(HE),OBJ:STAIRS;
4.3 STAIRS =LOC:(IN)APARTMENT,ATT:FLIGHTS;
4.4 FLIGHTS =NUM:TWO;
4.5 (POSS) PAT:HIS,OBJ:APARTMENT;
4.6 EQUIV:TEM (WHEN) [3.1],[4.1];

3) Protocol Analysis: In analyzing subjects' recall protocols of the clinical cases, a scoring procedure based on a specific set of rules for propositional matching (recall) and for propositional transformation (inference) developed by Frederiksen (1979,1981) was used. This procedure involves the use of the segmented recall protocols which are matched against the propositional analysis of the original stimulus texts.

The basic unit of comparison between the protocols and the stimulus texts is the **'triple'.** A triple consists of a concept-relation-concept unit. If the triple corresponds directly to the original message base then it is scored as a 'recall'. When the triple is the result of a transformation on the original message base,

then it is scored as an 'inference'. In view of the complexity of some propositions (e.g., 4.1 above), a recall and an inference can be scored on the same proposition.

In addition to recalls and inferences, there is one other catagory of response which is coded. This is the **'part-recall'.** A part-recall occurs when only a single concept or a relation-concept pair appears in the protocol. The recall is accurate, but does not consist of a codeable triple. This catagory was included in the analysis as a result of the subjects' use of point form notation in their protocols. Examples of each of these response types are given below.

Ten protocols were randomly selected to be analyzed by two expert coders as a check for consistency and reliability. The inter-rater reliabilities on these protocols ranged from 80.9 to 99.0 with a mean of 90.57% and a standard deviation of 4.81.

Protocol: No splenomegaly

Original Text:
1. THERE WAS NO SPLENOMEGALY

1.1 SPLENOMEGALY =NUM:NULL(NO),TNS:PAST;

Coding:
1.1 Recall

Protocol: He complained of dyspnea

Original Text:
3. HE ALSO COMPLAINED OF SOME SHORTNESS OF BREATH

3.1 COMPLAIN AGT:HE,THM:3.2,=TEM:___,
TNS:PAST;
3.2 BREATH =ATT:SHORTNESS,DEG:SOME;

Coding:
3.1 Recall
3.2 Inference (dyspnea- global term for shortness of breath)

Protocol: After he tried to climb two flights

Original Text:
4. WHEN HE TRIED TO CLIMB THE TWO FLIGHTS OF STAIRS
IN HIS APARTMENT

4.1	TRY	PAT:HE,GOAL:(TO)4.2,=TEM:4.6, TNS:PAST;
4.2	CLIMB	PAT:(HE),OBJ:STAIRS;
4.3	STAIRS	=LOC:(IN)APARTMENT,ATT:FLIGHTS;
4.4	FLIGHTS	=NUM:TWO;
4.5	(POSS)	PAT:HIS,OBJ:APARTMENT;
4.6	EQUIV:TEM (WHEN)	[3.1],[4.1];

Coding:
4.1 Recall and inference(change of when to after)
4.2 Part recall (only climb in protocol)
4.3 not coded
4.4 Recall
4.5 not coded
4.6 inference (change of EQUIV:WHEN to ORDER:AFTER)

4) Relevancy: The method for determining the relevancy of information in the clinical cases was developed by Patel et al. (1984) based on the work of Johnson (1970). This procedure involves a method of expert consensus.

Six (6) expert physicians were asked to read the clinical cases and to eliminate that text information which they did not consider to be highly relevant to the case. This was done so that the texts were reduced by 25-50% of their original lengths. The number of times that each unit of information was retained in the text provided a measure of the relevancy of the information. The inter-rater reliabilities for this procedure were 86% for the ABE case and 78% for the TA case.

This analysis of relevancy resulted in 29 propositions (32.95%) in the ABE text and 22 propositions (32.35%) in the TA text being designated as relevant to the diagnoses, while 59 propositions (67.05%) and 46 propositions (67.47%) were

designated as non-relevant in the ABE and TA cases, respectively.

5) Critical Information: The critical information in the cases was assessed in the same manner of consensus as the relevant information, only this time the experts were asked to determine the absolute minimum amount of information in each case required to reach an accurate diagnosis. The inter-rater reliabilities for this procedure were slightly higher than those for the relevant information.

The critical information was determined to consist of 8 propositions (9.09%) in the ABE case and 6 propositions (8.82%) in the TA case. The remaining 80 propositions (90.91%) in the ABE case and 62 propositions (91.18%) in the TA case were considered noncritical.

6) Statistical Analysis: A multivariate repeated measures analysis of variance was performed with the response types (recall, inference and part-recall) as the dependent variables and the level of expertise (expert or novice), case (ABE or TA), structure (structured or unstructured) and relevancy (high or low) as the independent variables. A second analysis was run identical to the first except that the relevancy factor was replaced by the critical information factor. A preliminary analysis was run which also included the text order variable. Text order was found to have no significant effects and was dropped from the analysis.

6) Diagnoses: The diagnoses of the subjects were grouped under three catagories: Accurate, Partially Accurate and Inaccurate. An accurate diagnosis was viewed as a complete diagnosis of the case (e.g., Acute Bacterial Endocarditis) while a partially accurate diagnosis was viewed as an incomplete, but accurate one (e.g.,

Endocarditis). An inaccurate diagnosis does not contain any aspect of the actual diagnosis (e.g., Rabies).

RESULTS AND DISCUSSION

The resultant MANOVA tables from the analyses are shown in Tables 1 and 2. All significant results involved within subjects factors or their interactions with the between subjects factors. None of the between subjects factors were significant on their own. Consistent with previous studies in the area, there was no significant main effect for level of expertise (L) on either analysis.

Relevant Information

The factor of response type (RS) was found to have a significant main effect ($F(2,27)=81.676$, $p<.0005$). This result reflects previous work in the area of text comprehension (e.g., Collins et al., 1980; Mandler and Johnson, 1977) which has shown that the direct recall of text information and the transformation of text information through inferencing reflect different mental processes in comprehension.

Previous studies in medicine by Patel et al. (1986) have consistently found more inferencing across all groups of subjects on pure comprehension tasks. In this study, where clinical cases representing patient problems are used, generating a solution to the problem (i.e. a diagnosis) is an integral part of the text

comprehension. On this task, all groups were found to recall more information than they inferred.

This finding is consistent with those in other areas such as comprehension of word arithmetic problems (Kintsch and Greeno, 1985) which found that, while there are general features in a comprehension process, the content of the specific comprehension strategies, nature of the knowledge structures involved and form of representation are both task and goal specific.

The fact that there is no level of expertise by structural form interaction (L x F) indicates that both experts and novices were similarly affected by the structural form factor. There is, however, a significant structural form by case type (F x C) interaction ($F(1,28)=9.334$, $p=.005$). At this level of analysis, there is virtually no difference in the amount of information which is contained in the protocols between the ABE case (15.89), and the TA case (15.87) based on their structured forms. However, in their unstructured forms there is an obvious decrement in the amount of information contained in the subjects' recall protocols for the ABE case (12.76), while there is less of a decrease in the TA case (13.88). This implies that disruptions in the text structure of a clinical case has more of an effect on the comprehension of an easier case than of a more difficult case. Also implied by this finding is that there may be some processing differences involved.

The importance of the relevance of information factor is reflected by the fact that there is a significant main effect for relevance of information ($F(1,28)=147.419$, $p<.0005$) and that all but one of the significant interactions found in this analysis involve the relevancy factor.

The fact that a significant relevance of information by response type

interaction (R x RS) was found ($F(2,27)=19.197$, $p<.0005$) implies that high and low relevance information are processed differently. This interaction is shown in Figure 1. Of more importance to any possible processing differences, however, are the differences within the two levels of relevancy. The proportion of information inferred in relation to information recalled is more in the case of nonrelevant information than for relevant. The findings of significant interactions between the level of expertise and relevance of information (L x R, ($F(1,28)=4.091$, $p=.053$)) and between the level of expertise, relevance of information and response type (L x R x RS, ($F(2,27)=3.566$, $p=.042$)) indicate more selective encoding of information on the part of the experts.

This relationship is represented by the significant L x R x RS interaction (Figure 2). While experts recall more high relevance information (31.36) than do novices (26.96), novices infer more on high relevance information (22.95) than do experts (19.85). Novices also recall (20.18) and infer (15.17) more low relevance information than do experts (17.14 and 11.67, respectively).

These differences indicate a basic processing difference between experts and novices. Not only do experts differentiate relevant from nonrelevant information, but they also process the two types of information differently.

The cases are obviously different in the way in which they were comprehended by both groups of subjects. Figure 3 shows the significant interaction of relevance of information by clinical case and response type (R x C x RS, ($F(2,27)=15.852$, $p<.0005$)). The aspect of interest here is that not only is the ability to selectively encode relevant information very different for the two cases, but that the processing of relevant and nonrelevant information is different both between and within cases. On the TA case, relevant and nonrelevant

information are neither differentiated to any extent nor are they processed differently.

On the ABE case, not only is relevant information selectively encoded, but the pattern of responses is different for high and low relevance information. This point brings us back to the initial question of the effect of text structure, in that it was found to interact with the clinical case factor.

If one views the response patterns across relevancy and text structure on the ABE case (Figure 4) the responses on this, the easier case, show a pattern which is consistent with the findings of Chase and Simon in the area of expert-novice differences in chess, while this pattern is different from those found in other studies in the area of medicine.

In the structured form of the ABE case, the experts recall far more relevant information (45.69) than do novices (35.78), who recall more of the nonrelevant information (21.82) than do experts (13.35). In the unstuctured form of the case these expert-novice differences in recall disappear for both the relevant (29.74 vs 31.03) and nonrelevant (10.59 vs 11.02) information.

In the more difficult, TA case (Figure 5) it appears that the difficulty of the case overrides any affects of the differences in the text structure. Less of the relevant information and more of the nonrelevant information is processed by both experts and novices across both text structures in the TA case compared to the ABE case. The experts do, however, recall more (25.00 and 24.99) and infer less (17.61 and 13.07) of the relevant information in the structured and unstructured forms of the TA case than do novices (recall 20.46 and 22.15; inference 21.02 and 18.18).

Critical Information

In analyzing the same data using the critical information factor, the significant effects of case type and structural form are eliminated, while the effect of level of expertise is enhanced. There is a significant level of expertise by response type interaction ($F(2,27) = 4.351$, $p=.023$) which was not present in the analysis of relevant information. This interaction clearly shows (Figure 6) the differences in the processing of information by the two experience groups. This difference is further represented by the level of expertise by critical information by response type interaction ($F(2,27) = 7.893$, $p=.002$) in Figure 7. The experts recalled more critical information (49.54%) than did the novices (29.24%). Novices, however, made inferences on more of the critical information (30.34%) than did the experts (21.54%). Novices also recalled and inferred more noncritical information than did the experts. Novices are not as selective as experts about the information they operate on and the two groups represent the cases differently in order to make a diagnosis.

Figures 8 and 9 show the overall results for critical information on the ABE case and the TA case, respectively. In viewing across the two cases there is an obvious consistency of the novices in their accessing and processing of critical and noncritical information irrespective of case or structure. The responses of the experts on the ABE case, however, still tend to resemble the findings of the Chase and Simon experiments. The experts responses in the unstuctured form of the ABE case much more closely resemble those of the novices than those in the structured form, where the experts are far superior to the novices in their recall of critical information. In the TA case the experts ability to access and recall

critical information is consistently high across structural forms.

Diagnoses

The diagnoses provided by the subjects are tabulated in Table 3. Among the novices, only two partially accurate diagnoses were provided, and both of these were provided by subjects who had read the structured form of the ABE case. These were 'bacterial infection-drug abuse' (subject N1) and 'endocarditis' (subject N6). All other diagnoses were incorrect and were invariably based on dramatic, but not necessarily relevant, text information that was more related to world knowledge than clinical expertise. For example, for the ABE case seven of the novices diagnosed 'rabies', presumably in response to the patient's "cat bite" story. In the TA case nine of the novices diagnosed 'psychological' or 'normal aging' problems in the "69 year old housewife".

Among the expert group, there were a total of four inaccurate diagnoses, one in each structural form of each case. Of major interest, and somewhat surprising, is the effect of the unstructured form of the text in the ABE case on the experts' comprehension. This effect was seen in the representation of the experts' protocols in Figure 5 and it also seems to have affected the quality of the diagnoses they provided.

For the structured form of the ABE case, four of the experts provided an accurate diagnosis, while for its unstructured form only one expert was totally accurate. The difficulty of experts in the ABE case was related to their inability to perceive the acuteness aspect of the case. This is represented by the fact that

three of the experts made incomplete diagnoses in the structured text and six of them in the unstructured text. This difficulty appears to have been exacerbated by the unstructured form of the text.

GENERAL DISCUSSION & IMPLICATIONS

The main findings of this study are the differential abilities of experts and novices to selectively encode and process relevant and critical clinical information.

The level of difficulty of the cases was found to affect the ability of all subjects to differentiate relevant from nonrelevant information and the way in which they process this information. The relevancy of information is clearly differentiated in an 'easier' clinical case and this high and low relevance information is processed in different ways. In a more difficult case, the relevancy of information is neither differentiated as efficiently nor are the levels of relevancy processed differently.

Access to critical information by experts, however, was not affected by case difficulty. Furthermore, this information was also used more efficiently in order to make a diagosis.

Case difficulty was also found to interact with the structural form of the cases. The effect of structural form in an easier clinical case replicates the findings in expert-novice studies in other areas, while a difficult clinical case is processed in a manner which mediates the effect of the structural form.

An analysis of the diagnoses provided by the subjects showed that the novices relied essentially on the text information in order to generate their diagnoses. The experts processed the information in the text at higher, more abstract levels, accessing their prior knowledge in order to generate their diagnoses. Further, the experts appeared to utilize two different processes in generating their diagnoses. One process was a form of pattern recognition and matching which proved effective in the structured form of the easy case. For the unstructured form of the easy case, where the "pattern" was disrupted, and for both forms of the difficult case, some other form of hypothesis generation was required.

These results show that there are expert-novice differences in clinical case repersentation and comprehension and that "atypicality", whether defined in terms of the difficulty or the structure of a case, does affect the way in which cases are processed. These differences, however, are not accessible through the use of global measures or methods of analysis. Finer levels of analysis such as propositional analysis are essential to the characterization of the effects of expertise in a field as complex as medicine.

In more specific terms, this study shows that the effect of the structural form of a clinical case on clinical case comprehension, and therefore the accuracy of diagnosis, may be related to the difficulty of that case. In a case which is easier to diagnose, the effects of structure replicate those effects which have been found in many other areas, including the classical studies of Chase and Simon (1973a,1973b) in chess. That is, in the structured form of the case, experts recall significantly more relevant information than do novices. This shows the importance of bottom level cues in the text comprehension process. In its unstructured form, however, these differences disappear. This argues for the role

of top level control in the process of comprehension. Thus both the top-down and bottom-up processes appear to be necessary.

However, all of the stimuli used by Chase and Simon are analogous only to the 'easier cases' used in the medical context. That is, they did not include any stimuli which are comparable to the difficult clinical case situation of the present study. The recall protocols provided for difficult cases differ from those for easy cases and do not replicate Chase and Simon. The protocols do not differ significantly across the structural forms of the difficult case. In a more difficult clinical case, the effect of the structural form appears to be mediated by the fact that the processing of the case information is different from that in an easier case.

Being a difficult case, the information relevant to the diagnosis does not exist as a specific inter-related unit of information. Furthermore, the text deals with signs and symptoms of a patient whose disease process is difficult to interpret in the form of a causal chain of events. However, the fact that the experts had no difficulty in diagnosing the case means that a set of rules or procedures critical to making a diagnosis must exist independent of any disease-specific information and can be applied in a rare or difficult case (eg., the TA case).

In clinical case comprehension, which implicitly includes the problem solving task of generating a diagnosis, the 'recall' of relevant and critical information is most important to generating an accurate diagnosis. This is in contrast to studies of pure comprehension of medical text information where it was found that the ability to make 'inferences' on relevant information is important both for comprehension and for differentiating experts and novices.

In the case of the Acute Bacterial Endocarditis text, there is a set of events which are linked in a causal chain that leads directly to the diagnosis. This pattern of temporally ordered events is disrupted by the unstructured form of the case. This is supported both by the different patterns of responses provided by experts on the two structural forms of the same case and by the quality of their diagnoses.

In the more difficult case, Temporal Arteritis, relevant and nonrelevant information are not clearly differentiated, but the experts are able to access critical information and their diagnoses are more accurate across the structural forms of the case. This implies a different, more rigorous generation of the diagnosis.

In terms of expert-novice differences, expertise in medicine is not unlike expertise in other areas. Experts in medicine are more efficient at selectively encoding and using relevant and critical information than are novices. As has been found in other areas, novices, in solving the problem of generating a diagnosis, perform a perceptual task utilizing mostly text information in reaching their diagnosis. Experts utilize selective text information at a higher, more abstract level, and based on their prior knowledge, they generate a diagnosis either through a process of pattern matching or through some form of hypothesis testing.

Many researchers in both education (e.g., Glaser, 1976; Larkin et al., 1980) and medicine (e.g., Feltovich, 1983) have expressed the importance of discovering those characteristics which differentiate experts and novices so as to better be able to facilitate the transition from the lower levels of performance to the higher levels.

The methodology required for this research has been well established in the

work of many cognitive psychologists, in many subject areas. In medicine, however, studies of expert-novice differences have not been able to fully satisfy the logical and intuitive expectations of either researchers or physicians. This may well be due to the fact that the field of medicine is by far the most difficult and complex subject area yet investigated by psychological researchers. A complex content area requires complex research techniques and analytical methods.

As the analysis of the comprehension and problem solving aspects of medicine become more complex and fine-grained the details of what it is that an expert physician does, will become more clearly defined. Future research must continue to look at medical expertise by focusing on ever more minute details of clinical case comprehension and clinical competence, but researchers must also begin to view their work in terms of the real world.

The greatest challenge to researchers is the fact that physicians, as a group, are very pragmatic. Empirical findings are only as good as their generalizability to everyday practice or teaching. The findings from laboratory and other controlled settings must be tested and verified under actual clinical conditions.

These series of studies are the beginnings of much work to come which will add to our knowledge of what it means to "understand". Such knowledge is necessary if we are to build sound models of instruction, which can be rigidly tested in the field.

The fields of medicine and medical education are presently entering a state of upheavel in which major tehnological changes in the everyday practice of medicine demand major changes in the education of practioners, both present and future. However, curricular and institutional changes can no longer be based purely on intuition. Empirically supported theories concerning the nature of the

components of clinical case comprehension and clinical competency are required in order to give purpose and direction to reforms.

The intuitions of medical practioners, as to what it is that they do in everyday practice and data collected from the behaviors of physicians are invaluable to the researcher and are an excellent starting point for our investigations. Empirical research based on these intuitions and findings can help to sort out the myth from the method and be instrumental in the development of both teaching and assessment techniques which can lead to a more fulfilling and rewarding educational experience for students, whether they be novices or experts, in the field of medicine.

REFERENCES

BARTLETT, F.C. (1932) Remembering. Cambridge University Press, Cambridge, England.

BOBROW, D.G. & NORMAN, D.A. (1975) Some principles of memory schemata. In Representation and Understanding. Bobrow & Collins (Eds.), Academic Press Inc., New York.

BORDLAND, R. & FLAMMER, A. (1985) Encoding and retrieval processes in memory for prose. Discourse Processes, **8**, 305-317.

BORDAGE, G. & ALLEN, T. (1982) The etiology of diagnostic errors: Process or content? Proceedings of the 21st Annual Conference on Research in Medical Education, American Association of Medical Colleges. Washington, D.C. 171-176.

CHASE, W.G. & SIMON, H.A. (1973a) Perception in chess. Cognitive Psychology, **4**, 55-81.

CHASE, W.G. & SIMON, H.A. (1973b) The mind's eye in chess. In Visual Information Processing, W.G. Chase (Ed.), Academic Press, London.

CHI, M.T.H., FELTOVICH, P.J. & GLASER, R. (1981) Categorization and representation of physics problems by experts and novices. Cognitive Science, **5**, 121-152.

CLAESSEN, H.F. & BOSHUIZEN, (1985) Recall of medical information by students and doctors. Medical Education, **19**, 61-67.

COLLINS, J., BROWN, J.S. & LARKIN, K.M. (1980) Inference in text understanding. In Theoretical Issues in Reading Comprehension. Spiro, Bruce & Brewer (Eds.),Lawrence Erlbaum & Associates, Hillsdale, N.J.

COUGHLIN, L.D. & PATEL, V.L. (1985) Adapting a paradigm from cognitive science to medical education: problems and possible solutions. Proceedings of the 24th Annual Conference on Research in Medical Education, American Association of Medical Colleges, Washington, D.C., 97-102.

DANKS, J.H. & GLUCKSBURG, S. (1980) Experimental psycholinguistics. Annual Review of Psychology, **31**, 391-417.

DEGROOT, A.D. (1965) Thought and Choice in Chess. Mouton Publishers, The Hague.

DILLINGER, M. (1984) Segmentation and Clause Analysis, Discourse Processing Laboratory, Report E3, McGill University, Department of Educational Psychology, Montreal, Quebec.

EGAN, D.E. & SCHWARTZ, B.J. (1979) Chunking in recall of symbolic drawings. Memory and Cognition, **7**, 149-158.

FELTOVICH, P.J. (1983) Expertise: Reorganizing and redefining knowledge use. Professions Education Researcher Notes, **4,5**.

FREDERIKSEN, C.H. (1975) Representing logical and semantic structure of knowledge acquired from discourse. Cognitive Psychology, **7**, 371-458.

FREDERIKSEN, C.H. (1979) Discourse comprehension and early reading. In Theory and Practice of Early Reading. L. Resnick & P. Weaver (Eds.). Lawrence Erlbaum & Associates, Hillsdale, N.J.

FREDERIKSEN, C.H. (1981) Inferences in preschool children's conversations: A cognitive perspective. In Language and Ethnography in Educational Settings, Green & Wallot (Eds.), Ablex, Norwood, N.J.

FREDERIKSEN, C.H. (1984) Frame related processes in comprehension and discourse communication. In I. Gopnik (Ed) Studies in Cognitive Science, Norwood, N.J.: Ablex

GENTNER, D.R. (1976) The structure and recall of narrative prose. Journal of Verbal Learning and Verbal Behavior, **15**, 411-418.

GLASER, R. (1976) Cognitive psychology and instructional design. In Cognition and Instruction. D. Klahr (Ed.), Hillsdale, N.J.

GROEN G.J. & PATEL V.L. (1985) Medical problem-solving: Some questionable assumptions. Medical Education, **19**, 95-100.

HALLIDAY, M. (1967) Notes on transitivity and theme in english. Part 2, Journal of Linguistics, **3**, 177-274.

HALLIDAY, M. & HASAN, R. (1976) Cohesion in English. Longmans, London.

HARRISON, T.R. (1980) Principles of Internal Medicine, McGraw-Hill Book Company, New York.

JOHNSON, R.E. (1970) Recall of prose as a function of the structural importance of the linguistic unit. Journal of Verbal Learning and Verbal Behavior, **9**, 12-20.

JOHNSON, R.E. (1982) Retrieval cues and the remembering of prose: A Review. In Discourse Processing, Flammer and Kintsch (Eds), North-Holland Publishing Co., Amsterdam

KINTSCH, W. (1974) The Representation of Meaning in Memory. Lawrence Erlbaum & Associates, Hillsdale, N.J.

KINTSCH, W. & GREENO J.G. (1985) Understanding and solving word arithmetic problems. Psychological Review, **92**, 109-129.

KINTSCH, W. & VAN DIJK, T. (1978) Toward a model of text comprehension and production. Psychological Review, **85**, 363-394.

KRULEE, G.K., FAIRWEATHER, P.G. & BERQUIST, S.R. (1979) Organizing factors in the comprehension and recall of connected discourse. Journal of Psycholinguistic Research, **8**, 141-163.

LARKIN, J.H., MCDERMOTT, J., SIMON, D.P. & SIMON, H.A. (1980) Models of competence in solving physics problems. Cognitive Science, **4**, 317-345.

MANDLER, J., & JOHNSON, N. (1977) Rememberance of things parsed: Story structure and recall. Cognitive Psychology, **9**, 111-151.

MUZZIN, L.J., NORMAN, G.R., JACOBY, L.L., FEIGHTNER, J.W., TUGWELL, P. & GUYATT, G.H. (1982) Manifestations of expertise in recall of clinical protocols. Proceedings of the 21st Annual Conference on Research in Medical Education, 163-168.

NORMAN, D.A. 7 RUMELHART, D.E. (EDS) (1975) Explorations in Cognition. Freeman Press, San Francisco.

NORMAN, G.R., JACOBY, L.L., FREIGHTNER, J.W. & CAMPBELL,G.J.M. (1979) Clinical experience and the structure of memory. Proceedings of the8th Annual Conference on Research in Medical Education, Washington, D.C. 214-218.

NORMAN, G.R., TUWELL, P., GEIGHTNER, J.W., MUZZIN, L.J. & JACOBY, L.L. (1985) Knowledge and clinical problem-solving. Medical Education, **19,5:** 344-356.

PATEL, V.L. & FREDERIKSEN, C.H. (1984a) Cognitive processes in comprehension and knowledge acquisition by medical students and physicians. In Tutorials in Problem-Based Learning, H.G. Schmidt & M.C. deVolder (Eds.), Assen, Holland: Van Gorcum.

PATEL, V.L. & FREDERIKSEN, C.H. (1984b). Role of Prior Knowledge in Expert-Novice Differences in Comprehension of Medical Texts. Report No. CME 84-5, McGill University: Montreal.

PATEL, V.L. & GROEN, G.J. (1986,in press) Knowledge based solution strategies in medical reasoning. Cognitive Science, **10,** 91-116.

PATEL, V.L., GROEN, G.J. & FREDERIKSEN, C.H. (1986,in press) Differences between students and physicians in memory for clinical cases. Medical Education, **20.**

PATEL, V.L., HOPINGKONG, H. & MARK, V. (1984) The role of prior knowledge in interpretation of medical texts by medical students and physicians. Proceedings of Research in Medical Education, Association of American Medical Colleges, Wasington D.C., 127-132.

PICHERT, J.W. & ANDERSON, R.C. (1977) Taking different perspectives on a story. Journal of Educational Psychology, **69,** 309-315.

REITMAN, J.S. (1976) Skilled perception in GO. Cognitive Psychology, **8,** 336-356.

TIERNEY, R.J. & MOSENTHAL, J. (1982) Discourse comprehension and production: Analyzing text structure and cohesion. In Reader Meets Author/ Bridging the

Gap: A psycholinguistic and sociolinguistic perspective, International Reading Association, Newark, DE.

WINOGRAD, T. (1972) Understanding Natural Language. Edinburgh University Press, Edinburgh.

WINOGRAD, T. (1983) Language as a Cognitive Process: Vol. 1 Syntax, Addison-Wesley Publishing Co. Inc., Reading, Mass.

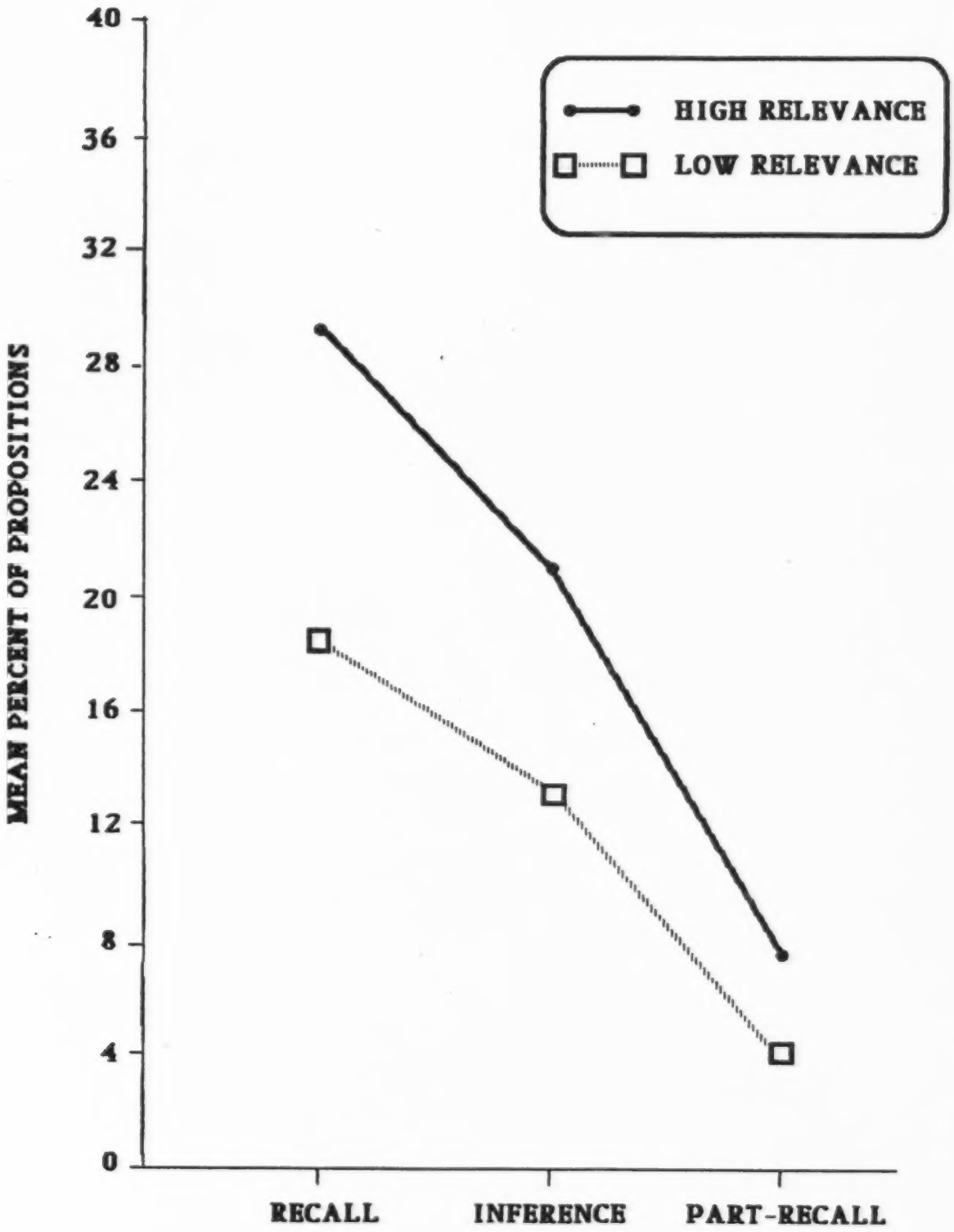

Figure 1 Mean Percent of Propositions by Relevance of Information and Response Type.

F(2,27) = 19.197, p< .0005

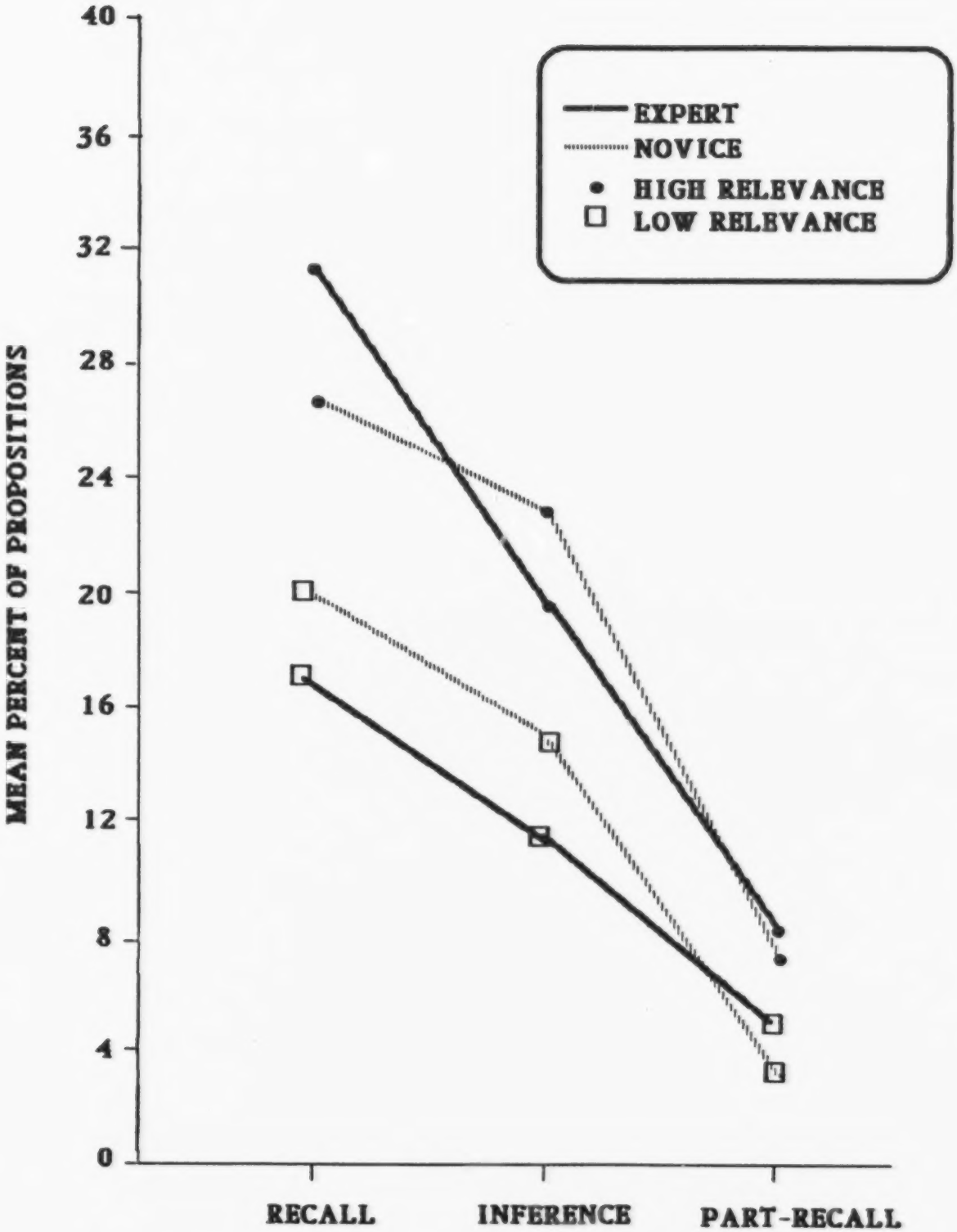

Figure 2 Mean Percent of Propositions by Level of Expertise, Relevance of Information and Response Type. F(2,27) = 3.566, p= .042

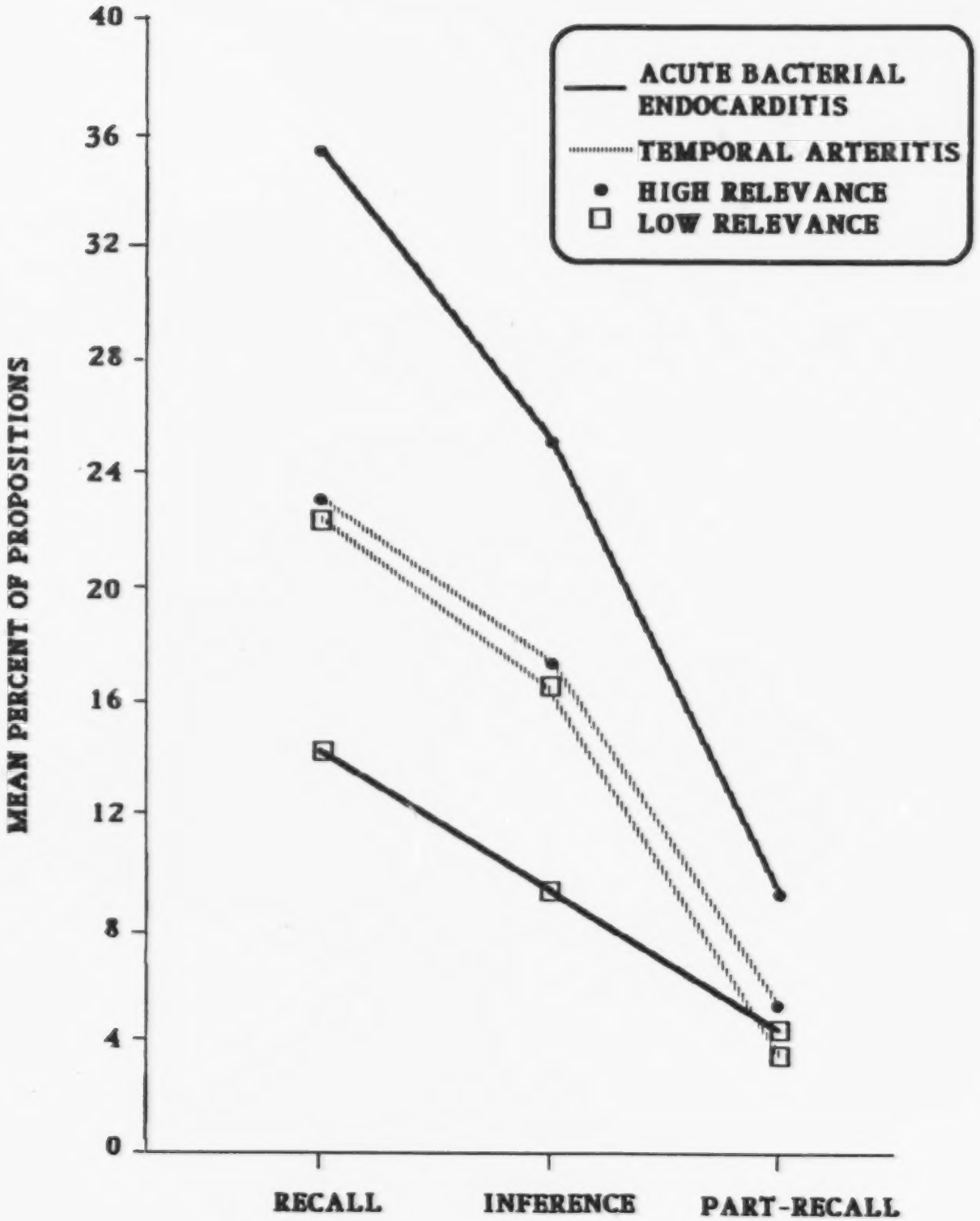

Figure 3 Mean Percent of Propositions by Clinical Case, Relevance of Information and Response Type. $F(2,27) = 15.852$, $p < .0005$

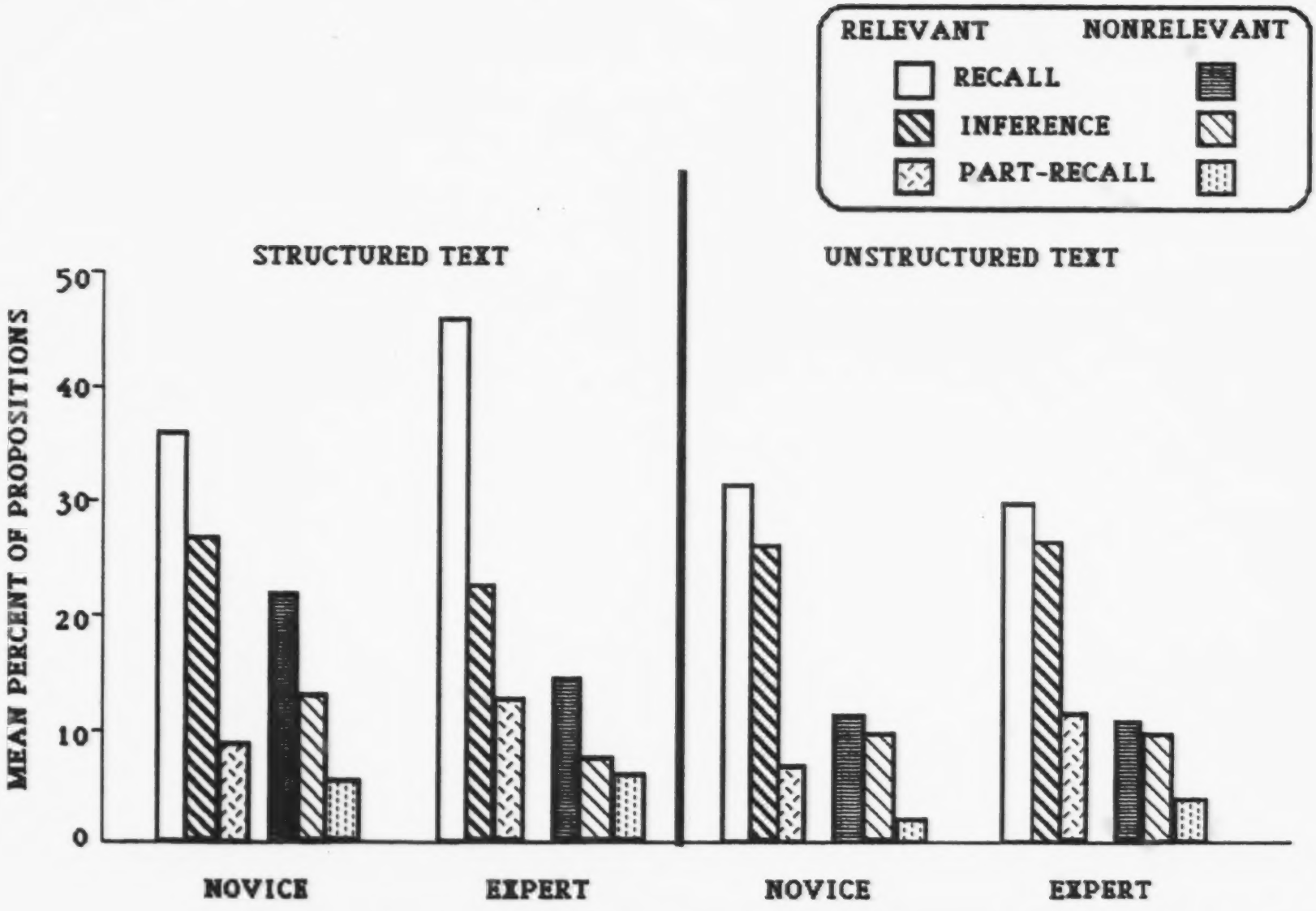

Figure 4 Mean Percent of Propositions by Relevance of Information, Level of Expertise, Structural Form and Response Type on the Acute Bacterial Endocarditis Case.

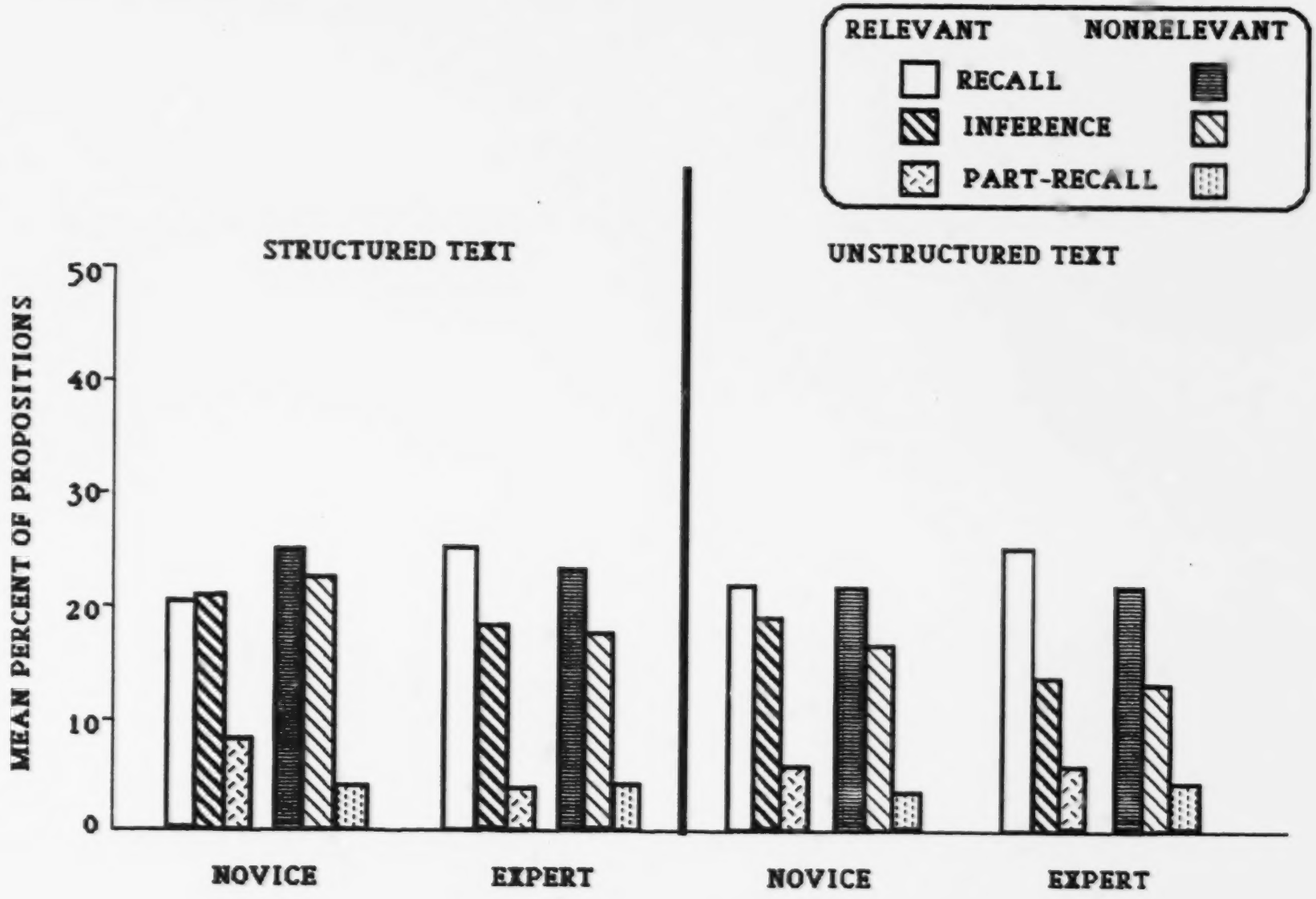

Figure 5 Mean Percent of Propositions by Relevance of Information, Level of Expertise, Structural Form and Response Type on the Temporal Arteritis Case.

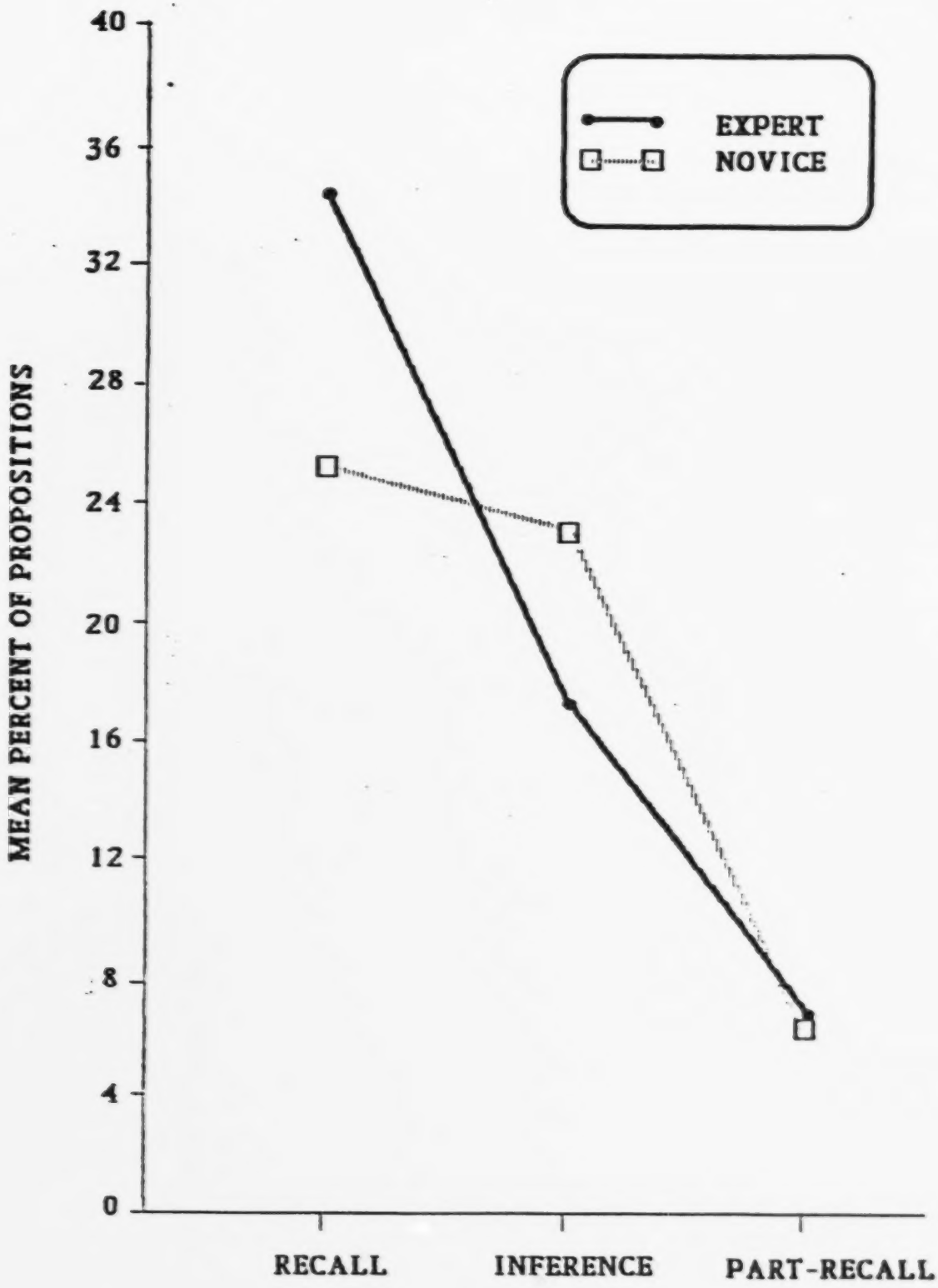

Figure 6 Mean Percent of Propositions by Level of Expertise and Response Type
F(2,27) = 4.351, p= .023

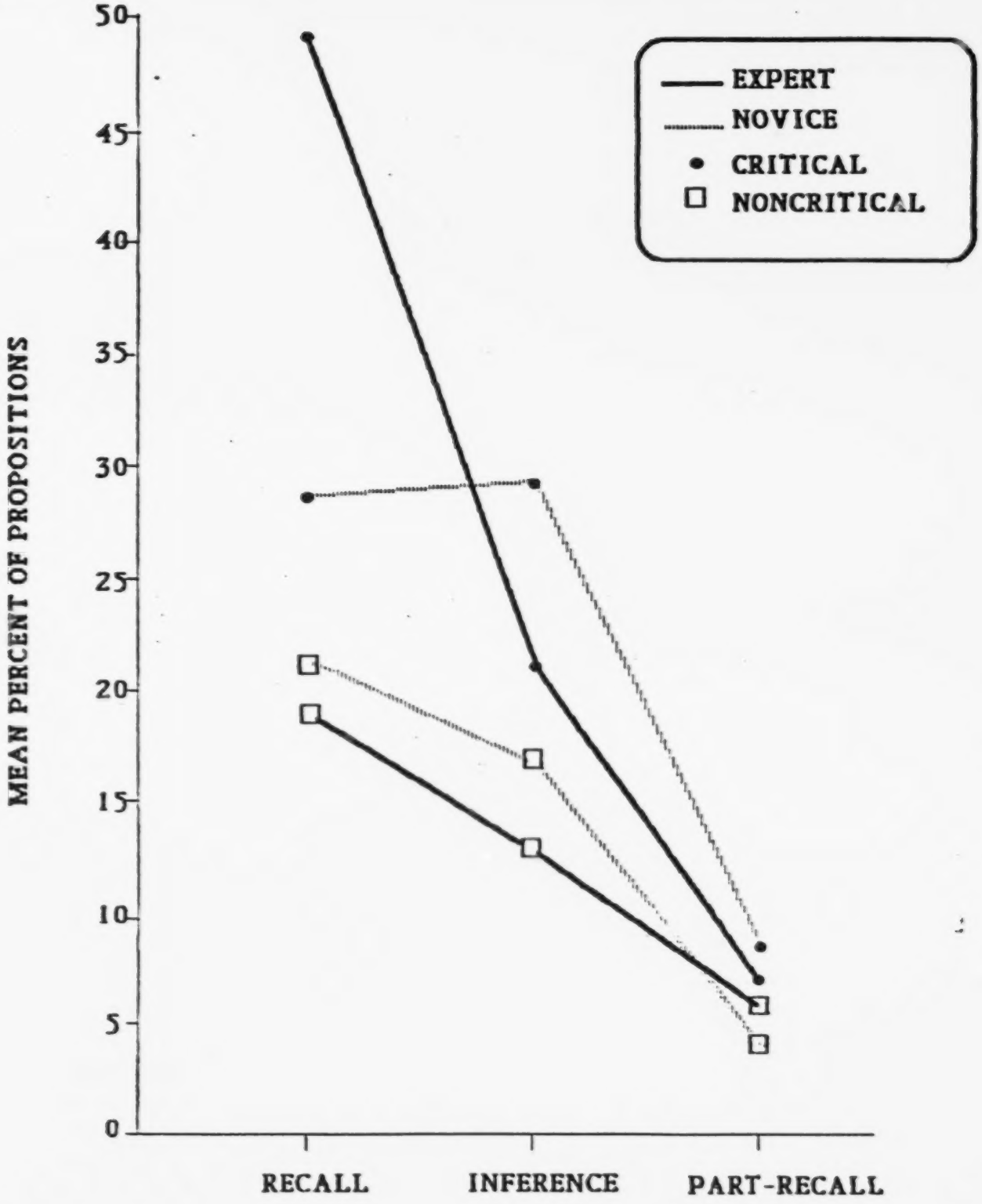

Figure 7 Mean Percent of Propositions by Level of Expertise, Critical Information and Response Type. $F(2,27) = 7.893$, $p = .002$

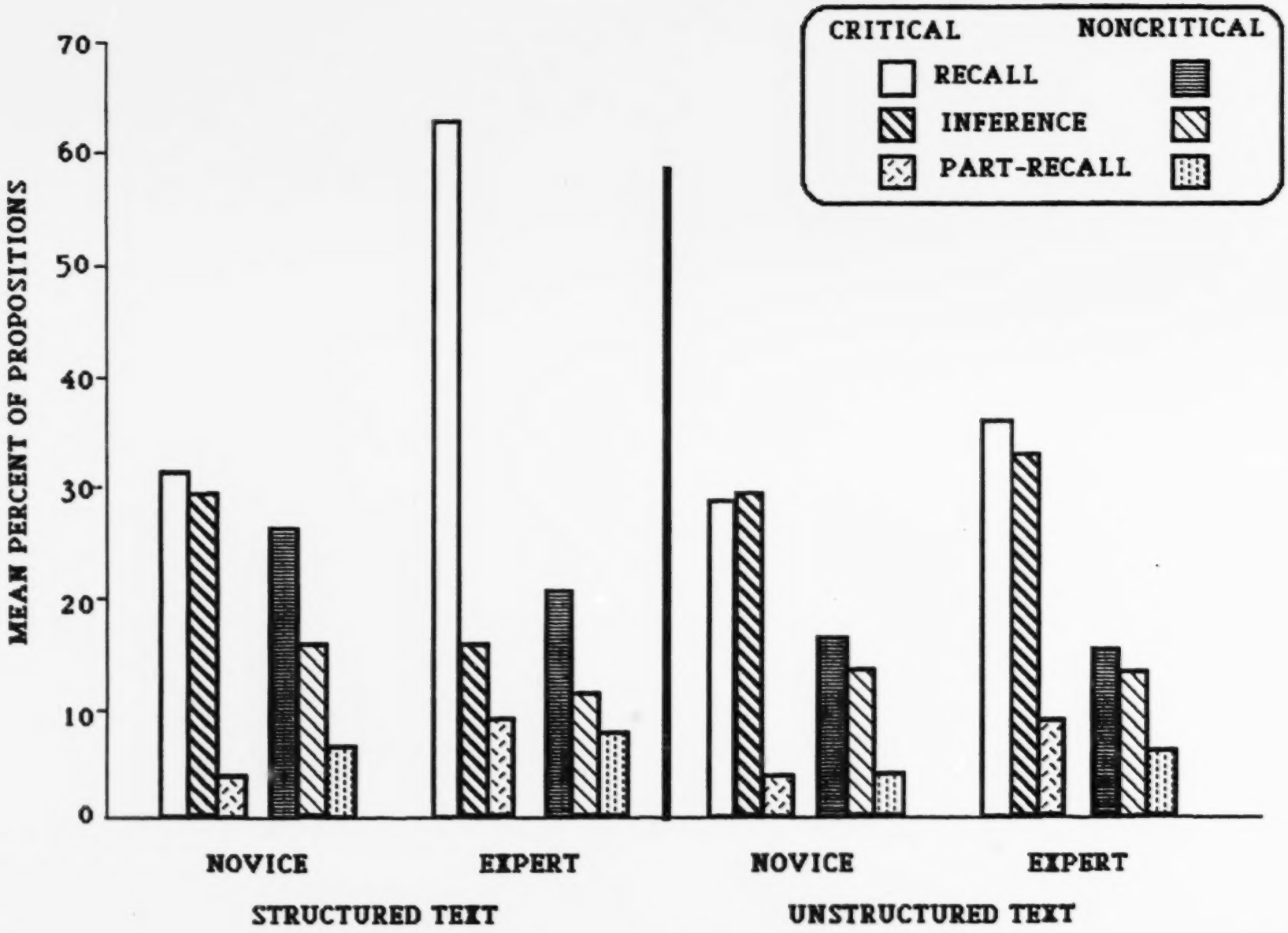

Figure 8. Mean Percent of Propositions by Critical Information, Level of Expertise, Structural Form and Response Type on the Acute Bacterial Endocarditis Case.

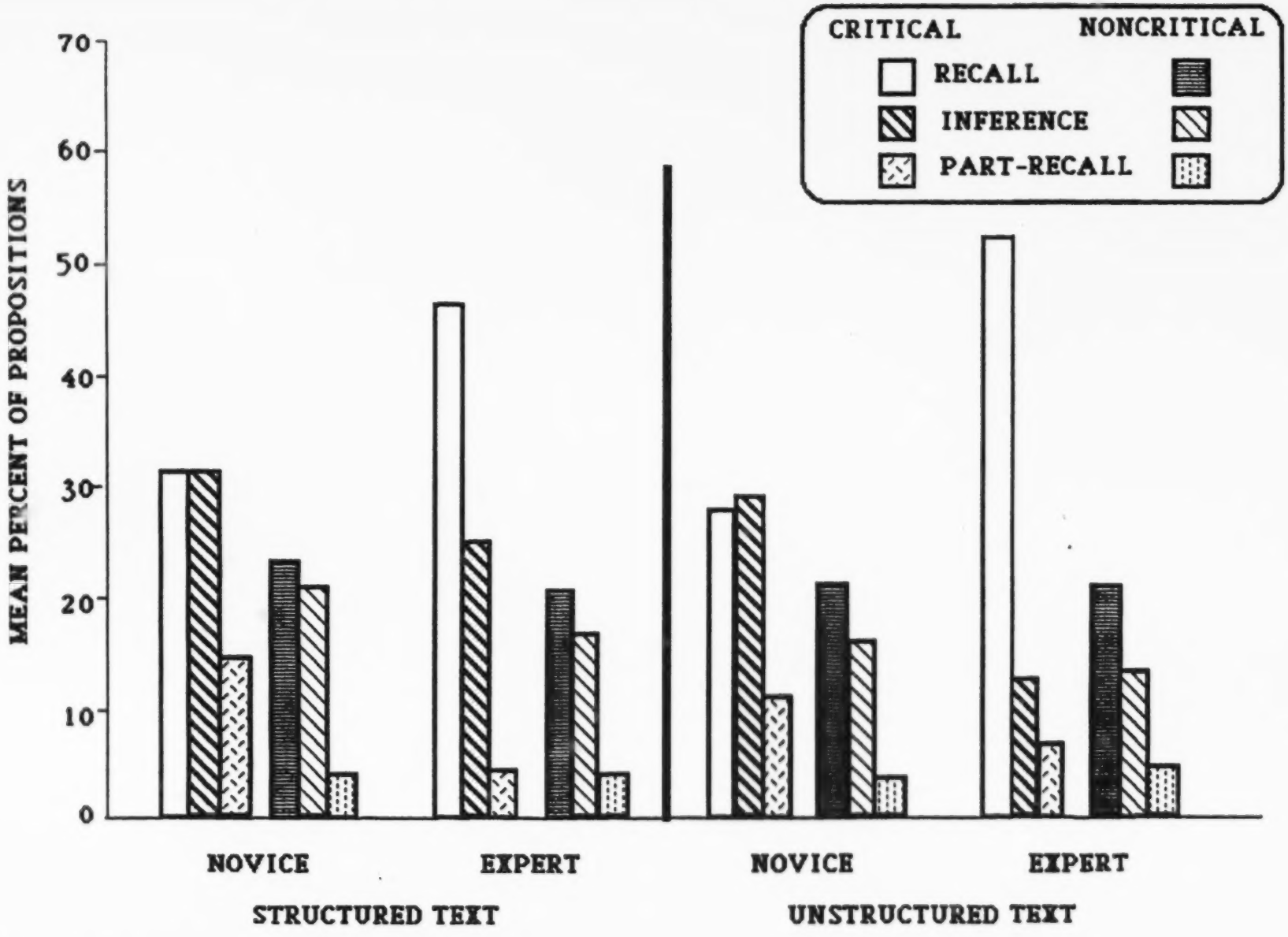

Figure 9. Mean Percent of Propositions by Critical Information, Level of Expertise, Structural Form and Response Type on the Temporal Arteritis Case.

TABLE 1
Results of the Repeated Measures Multivariate Analysis of Variance (Relevant Information)

SOURCE	F	df	SIGNIFICANCE
Constant	492.28	1,28	.000
Between subjects			
Level of expertise(L)	< 1		
Structural form(F)	< 1		
L x F	< 1		
Within subjects			
Relevance(R)	147.419	1,28	.000 ***
Case type(C)	2.847	1,28	.103
Response type(RS)	81.676	2,27	.000 ***
R x C	129.573	1,28	.000 ***
R x RS	19.197	2,27	.000 ***
C x RS	1.367	2,27	.272
R x C x RS	15.852	2,27	.000 ***
Interactions of Between and Within Factors			
L x R	4.091	1,28	.053 *
L x C	< 1		
L x RS	1.667	2,27	.208
F x R	< 1		
F x C	9.334	1,28	.005 **
F x RS	1.962	2,27	.160
L x F x R	1.028	1,28	.319
L x F x C	<1		
L x F x RS	<1		
L x R x C	2.116	1,28	.157
L x R x RS	3.566	2,27	.042 *
L x C x RS	<1		
F x R x C	<1		
F x R x RS	<1		
F x C x RS	1.518	2,27	.237
L x F x R x C	3.604	1,28	.068
L x F x R x RS	<1		
L x F x C x RS	<1		
L x R x C x RS	<1		
F x R x C x RS	<1		
L x F x R x C x RS	1.419	2,27	.259

TABLE 2
Results of the Repeated Measures Multivariate Analysis of Variance (Critical Information)

SOURCE	F	df	SIGNIFICANCE
Constant	505.97	1,28	.000
Between subjects			
Level of expertise(L)	< 1		
Structural form(F)	< 1		
L x F	< 1		
Within subjects			
Relevance(R)	74.239	1,28	.000 ***
Case type(C)	< 1		
Response type(RS)	66.424	2,27	.000 ***
R x C	< 1		
R x RS	12.063	2,27	.000 ***
C x RS	< 1		
R x C x RS	2.283	2,27	.121
Interactions of Between and Within Factors			
L x R	3.995	1,28	.055 *
L x C	1.536	1,28	.226
L x RS	4.351	2,27	.023 *
F x R	< 1		
F x C	3.905	1,28	.058
F x RS	2.535	2,27	.098
L x F x R	< 1		
L x F x C	< 1		
L x F x RS	1.471	2,27	.248
L x R x C	2.063	1,28	.162
L x R x RS	7.893	2,27	.002**
L x C x RS	< 1		
F x R x C	< 1		
F x R x RS	1.116	2,27	.342
F x C x RS	1.520	2,27	.237
L x F x R x C	< 1		
L x F x R x RS	2.444	2,27	.106
L x F x C x RS	< 1		
L x R x C x RS	< 1		
F x R x C x RS	< 1		
L x F x R x C x RS	< 1		

TABLE 3
Diagnoses Provided by Experts and Novices

	ACUTE BACTERIAL ENDOCARDITIS (1-8 TYPICAL FORM) (9-16 ATYPICAL FORM)			TEMPORAL ARTERITIS (1-8 ATYPICAL FORM) (9-16 TYPICAL FORM)		
	ACCURATE	PARTIAL	INACCURATE	ACCURATE	PARTIAL	INACCURATE
NOVICES						
1		X				X
2			X			X
3			X			X
4			X			X
5			X			X
6		X				X
7			X			X
8			X			X
9			X			X
10			X			X
11			X			X
12			X			X
13			X			X
14			X			X
15			X			X
16			X			X
EXPERTS						
1	X			X		
2	X			X		
3	X			X		
4		X		X		
5		X				X
6			X	X		
7	X				X	
8		X		X		
9		X				X
10		X		X		
11		X		X		
12	X			X		
13		X		X		
14			X	X		
15		X		X		
16		X		X		